

Playing games in an uncertain world

Manoj Gopalkrishnan*, Girish Varma†

School of Technology and Computer Science,
Tata Institute of Fundamental Research, Mumbai, India.

February 20, 2014

Abstract

Traditional game theory assumes that the players in the game are aware of the rules of the game. However, in practice, often the players are unaware or have only partial knowledge about the game they are playing. They may also have knowledge that other players have only partial knowledge of the game they are playing, which they can try to exploit. We present a novel mathematical formulation of such games. We make use of Kripke semantics, which are a way to keep track of what different players know and do not know about the world. We propose a notion of equilibrium for such games, and show that equilibrium always exists.

1 Introduction

Recall the Monty Hall problem. A games show host shows a contestant three doors, and tells the contestant that there are prizes behind one of the doors, and no prize behind the other two. The contestant is asked to pick a door. The games show host opens one of the other doors, revealing that it has no prize. He asks the contestant if they want to switch their choice, or stick with their original choice.

It is well known that most people intuitively feel it makes no difference if they switch or not, whereas the best strategy is to indeed switch. The standard explanation given for this apparently suboptimal human behavior has been that human beings are not good with Bayes' theorem, and probabilistic reasoning. This may well be true. However, there is another possible explanation. Perhaps the suboptimal behavior has to do with the contestant not being completely clear on what the rules of the game are.

Perhaps the contestant is not sure about the following questions: Did the show host know that the door he threw open did not have anything behind it? Did he use that fact

*manojg@tifr.res.in

†girishrv@tifr.res.in

while making a choice of that door? Or did he randomly select a door, and then reveal it to have no prize behind it? What is the incentive for the games show host: help the contestant win a prize, make sure the contestant wins as less as possible, or entertain the audience?

Ambiguity in the rules of the game being played is very common in real life. The payoff matrix for a game often depends on information that the players may possess only partially. To complicate things, players may have partial information about about the information available to other players. This can lead to higher-order strategic reasoning based on what other players know and do not know.

Our contributions are:

- Kripke semantics is a well-developed field concerned with reasoning about knowledge. By combining Kripke semantics with game theory, we come up with a mathematical formulation we call **Kripke games** (Definition 3.3). Our formulation allows a precise, direct, and granular treatment of the effect of epistemic considerations on strategy.
- We propose a notion of **equilibrium** for Kripke games (Definition 3.5). Our notion assumes that every player supposes the worst about her missing knowledge, and that this pessimism among players is common knowledge. This definition, within its limitations, has the merit that it succinctly captures the effect of arbitrary levels of epistemic reasoning among the players. We prove in Theorem 3.2 that every Kripke game has an equilibrium.
- We introduce the notion of **min-games** (Definition 4.1). The idea is that the players are simultaneously involved in multiple games, and each player’s payoff is the minimum across all these games. In Theorem 5.1, we prove that every Kripke game with pessimistic players reduces to a min-game, and in Theorem 4.2, we prove that every min-game has an equilibrium.

Apart from their mathematical role in the above theorems, min-games appear to be interesting objects of study in their own right. One-player min-games correspond to optimization tasks where the payoff is the minimum among various payoffs. This setting is closely related to multi-objective optimization [4], which is an active research area. Thus min-games may be viewed as a natural extension of multi-objective optimization to the context of games.

2 An Example

In traditional two-player zero-sum games [5], one considers a game between two players, *Row* and *Column*. Each player has a choice of two pure strategies, strategy 1 and strategy 2. Their payoffs are described by a single 2×2 **payoff matrix** with rows indexed by

strategies of *Row*, and columns indexed by strategies of *Column*. *Row* tries to maximize payoff, and *Column* tries to minimize it.

What if the players don't know exactly what game they are playing? For concreteness, consider three possible worlds, call them world 1, world 2, and world 3, with three different payoff matrices, as follows:

- In world 1, the payoff matrix is given by $\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$.
- In world 2, the payoff matrix is given by $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$.
- In world 3, the payoff matrix is given by $\begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$.

If both players know the “true world” (or a probability distribution over the possible worlds) then this reduces to a two-player zero-sum game in the traditional sense of von Neumann and Morgenstern, and can be solved in the usual way. However, suppose the players don't even know a probability distribution over the possible worlds. In addition, though the players don't know which world is the true world, perhaps *Row* knows that the true world is either 1 or 2, whereas *Column* knows that the true world is either 2 or 3, and *Row* has some idea about what *Column* may know, and vice versa. This sort of strategic thinking is what we are interested in exploring. We will capture this through the device of **Kripke semantics** [3] on the set of worlds.

We're going to introduce two equivalence relations \equiv_r and \equiv_c , corresponding to the row and column player respectively, on the set $\{1, 2, 3\}$ of worlds. The idea is that all \equiv_r worlds look identical to *Row*, and similarly all \equiv_c worlds look identical to *Column*. For concreteness, we define \equiv_r and \equiv_c as the smallest equivalence relations containing the relations $1 \equiv_r 2$ and $2 \equiv_c 3$. In detail:

- The equivalence relations themselves are assumed to be common knowledge between *Row* and *Column*. So both *Row* and *Column* know $1 \equiv_r 2$ and $2 \equiv_c 3$, and both know that both know it, and so on.
- If the true world is world 2, then *Row* knows that the true world is either world 1 or world 2, since $1 \equiv_r 2$, whereas *Column* knows that the true world is either world 2 or world 3, since $2 \equiv_c 3$.
- If the true world is world 1, then *Row* knows that the true world is either 1 or 2, since $1 \equiv_r 2$. *Column* knows that the true world is world 1 since $w \equiv_c 1$ implies $w = 1$. Further, *Column* knows that *Row* knows that the true world is either 1 or 2. And so on.

So if the true world is world 3, *Row* knows that the true world is world 3. So *Row* could simply play the optimal strategy for Game 3. However, *Row* knows that *Column* does not know if the true world is world 2 or world 3! Possibly *Row* can exploit this knowledge to do even better.

A computation shows that under the notion of equilibrium introduced in Definition 3.5, the optimal strategy for *Row* is to play either row with probability 1/2 if the world is either 1 or 2, and to play the second row if the world is 3. The optimal strategy for *Column* is to play either column with probability 1/2 if the world is 1. If the world is 2 or 3, then *Column* plays the first column with probability 0.6. The corresponding expected payoffs in the three worlds are 0, 0, 4/5 to *Row*, and 0, -4/5, -4/5 to *Column* respectively.

3 Kripke Games

We first recall the well-known notions of finite (non-zero sum, non-cooperative) games, taking the opportunity to establish notation.

Definition 3.1 (Finite Game). Let P be a finite set. A **finite game with players P** is described by: for each player $p \in P$, a finite set S_p of *strategies*, and a *payoff* $u_p : \prod_{q \in P} S_q \rightarrow \mathbb{R}$.

Definition 3.2 (Mixed strategy). Fix a finite game with players P .

1. For each player $p \in P$, a **mixed strategy** is a probability distribution σ_p over S_p .
2. A **play** is a probability distribution $\sigma = \prod_{p \in P} \sigma_p$ on $\prod_{p \in P} S_p$, where each σ_p is a mixed strategy for player p .
3. A **play instance** is a random variable $X = (X_p)_{p \in P}$ such that $\text{law}(X)$ is a play, i.e., each X_p takes values in S_p , and $\{X_p\}_{p \in P}$ are all independent.
4. The **payoff** $\text{pay}_p(\sigma)$ to player p from a play σ is the expected payoff $\mathbb{E}[u_p(X)]$ where X is a play instance with law σ .

We are now ready to introduce Kripke games.

Definition 3.3 (Kripke Game). Let P and W be finite sets. A **Kripke game with players P on worlds W** is described by: for each player $p \in P$,

- A finite set S_p of *strategies*,
- An equivalence relation \equiv_p on the set W of worlds,
- For each world $w \in W$, a *payoff* $u_p^w : \prod_{q \in P} S_q \rightarrow \mathbb{R}$.

Remark 3.1. Note that the set S_p of strategies of player p is the same in all worlds. There is no loss of generality in this assumption since, if a player is allowed different strategies on different worlds, we can always reduce to this case by the following trick. We take the union of all possible strategies and declare that to be the set S_p . We extend the payoff matrix, by making the payoff $-A$ if the player plays a previously non-existent strategy in a world, where A is a real number sufficiently large to ensure that the player will never want to play that strategy.

Notation 3.3.1. $[W]_p$ will denote the set of \equiv_p -equivalence classes of W . If $w \in W$ then $[w]_p \in [W]_p$ will denote the \equiv_p -equivalence class containing w . That is, $[w]_p = \{v \in W \mid v \equiv_p w\}$.

Definition 3.4 (Strategy). Fix a Kripke game with players P on worlds W .

1. A **pure strategy** s_p for player p is a map from $[W]_p$ to S_p .
2. A **mixed strategy** σ_p for player p is a probability distribution over the pure strategies $S_p^{[W]_p}$ of player p .
3. A **play** is a probability distribution $\sigma = \prod_{p \in P} \sigma_p$, where each σ_p is a mixed strategy.
4. A **play instance** is a random variable $X = (X_p)_{p \in P}$ such that $\text{law}(X)$ is a play, i.e., each X_p takes values in $S_p^{[W]_p}$, and $\{X_p\}_{p \in P}$ are all independent.
5. The **specialization** $X(w)$ of a play instance $X = (X_p)_{p \in P}$ to the world w is the random variable $(X_p([w]_p))_{p \in P}$ taking values in $\prod_{p \in P} S_p$. We will view $X(w)$ as a play instance on the finite game in the world w .
6. The **payoff** $\text{pay}_p^w(\sigma)$ for player p in world w from play σ is the minimum over all worlds $v \in [w]_p$ of the expected payoff $\mathbb{E}[u_p^v(X(v))]$, where X is a play instance with law σ . That is:

$$\text{pay}_p^w(\sigma) := \min_{v \in [w]_p} \mathbb{E}[u_p^v(X(v))].$$

The definition of the payoff pay_p^w requires further explanation. Note that if $v \equiv_p w$ then in fact $\text{pay}_p^v = \text{pay}_p^w$. The idea is that if the true world is w , then player p only knows that the true world belongs to the equivalence class $[w]_p$. So player p takes a worst-case attitude, assuming that her payoff will be the worst payoff over all worlds $v \in [w]_p$.

Notation 3.4.1. Fix a Kripke game with players P . We will denote by Σ_p the set of mixed strategies for player p , and by Σ the set $\prod_{p \in P} \Sigma_p$ of all plays. If $\sigma = \prod_{p \in P} \sigma_p$ is a play, and σ'_q is a mixed strategy for player q then $\sigma - \sigma_q + \sigma'_q$ will denote the play $\sigma'_q \times \prod_{p \in P \setminus \{q\}} \sigma_p$ obtained from σ by replacing the mixed strategy σ_q of player q by the mixed strategy σ'_q .

Definition 3.5 (Equilibrium in a Kripke Game). Fix a Kripke game with players P on worlds W .

1. A player $p \in P$ **tolerates** a play σ iff for all worlds $w \in W$, for all mixed strategies $\sigma'_p \in \Sigma_p$, the payoff $\text{pay}_p^w(\sigma) \geq \text{pay}_p^w(\sigma - \sigma_p + \sigma'_p)$.
2. A play σ is an **equilibrium** iff every player tolerates σ .

Note that our notion of equilibrium assumes that each player is pessimistic/ risk-averse, and that this is common knowledge. What is the psychological validity of our assumption? There is some evidence in the psychology literature in the work of Kahneman and Tversky [2] that human beings are indeed risk-averse in decision making. It is an important question, and possibly one that lends itself to empirical investigation, whether and in what settings human beings treat such risk-averseness among all human players as common knowledge.

We emphasize that Kripke games should admit various other notions of equilibrium, for example, inspired by ideas from bounded rationality, or repeated games, which we will not explore further in this paper. Our main theorem is the following.

Theorem 3.2. Every Kripke game has an equilibrium play.

In Section 4, we define **min-games**, and show that they permit a notion of equilibrium. In Section 5, we show that Kripke games reduce to min-games, so that equilibrium plays in min-games allow us to obtain equilibrium plays in Kripke games.

4 Min-Games

Example 4.1. Consider two zero-sum finite games between *Row* and *Column*, described by the payoff matrices $\begin{pmatrix} -1 & 2 \\ 1 & -3 \end{pmatrix}$ and $\begin{pmatrix} 2 & -4 \\ 0 & 5 \end{pmatrix}$, where *Row* wants to maximize payoff, and *Column* wants to minimize. Suppose *Column* plays the first column. If *Row* plays the first row, her payoff is -1 from the first game, and 2 from the second game. If she plays the second row, her payoff is 1 from the first game, and 0 from the second game.

In a min-game, we assume that the payoff to a player is the worst-case of her payoff from all the games. So *Row*'s min-game payoff in response to *Column* playing the first column is -1 if she plays the first row, and 0 if she plays the second row. Her best response to *Column* playing the first column is thus to play the second row.

Column's payoffs to the play (row 2, column 1) are 1 and 0 . Since *Column* is trying to minimize, her worst-case payoff is the maximum of these two values, i.e., 1 . Note that in this play, *Row*'s min-game payoff is coming from the second game, and *Column*'s min-game payoff is coming from the first game. So it is possible for different players to get their min-game payoff values from different games. We now formalize these ideas, and introduce a notion of equilibrium for min-games.

Definition 4.1 (Min-Game). Fix a finite set P and a positive integer $k \in \mathbb{Z}_{\geq 1}$. A **min- k -game** with players P is described by: for each player $p \in P$, a set S_p of strategies, and payoffs $u_p^1, u_p^2, \dots, u_p^k : \prod_{q \in P} S_q \rightarrow \mathbb{R}$.

Remark 4.1. The intuitive idea is that there are k parallel games going on. We are assuming that the set of strategies for each player is the same in all the k games she is playing. This causes no loss of generality, by an argument parallel to the argument offered in Remark 3.1.

The definitions of mixed strategy, play, and play instance carry over to min-games unchanged from the setting of finite games in Definition 3.2. The key difference is in the definition of payoff.

Definition 4.2. Fix a min- k -game with players P .

1. The **payoff** $\text{pay}_p(\sigma)$ to player p from a play σ in a min- k -game is defined by

$$\text{pay}_p(\sigma) := \min \left\{ \mathbb{E}[u_p^1(X)], \mathbb{E}[u_p^2(X)], \dots, \mathbb{E}[u_p^k(X)] \right\}$$

where X is a play instance with law σ .

2. Player p **tolerates** play σ iff for all mixed strategies σ'_p for player p , the payoff

$$\text{pay}_p(\sigma - \sigma_p + \sigma'_p) \leq \text{pay}_p(\sigma).$$

3. An **equilibrium** is a play tolerated by every player.

We now state the main theorem about min-games, which shows existence of equilibria.

Theorem 4.2. Every min-game has an equilibrium play.

Remark 4.3. Consider a one-player min-game where the player has two strategies s_1 and s_2 , and the payoff functions are $u^1(s_1) = 1, u^1(s_2) = 2, u^2(s_1) = 3, u^2(s_2) = 0$. Then $\text{pay}(ps_1 + (1-p)s_2)$ is the minimum of $p + 2(1-p)$ and $3p$. This minimum is achieved at $p^* = 1/2$ and has value $3/2$.

In particular, note that this game can not be represented by a one-player finite game, because every one-player finite game has a linear payoff function, whereas for this game the payoff function is not linear.

We will give two proofs for Theorem 4.2. The first proof will employ Kakutani's fixed point theorem in a more general way than Nash's proof for existence of equilibrium in finite games. Recall the statement of Kakutani's fixed point theorem from [6].

Theorem 4.4 (Kakutani's fixed point theorem). Let S be a non-empty, compact, and convex subset of some Euclidean space \mathbb{R}^n . Let $B : S \rightarrow 2^S$ be a set-valued function on S with a closed graph and the property that $B(x)$ is non-empty and convex for all $x \in S$. Then B has a fixed point.

4.1 The first proof

I for Theorem 4.2. Fix a min- k -game with players P . For each player $p \in P$, let Σ_p denote the set of mixed strategies of player p . Let $\Sigma = \prod_{p \in P} \Sigma_p$ be the set of all plays. We will define an *improvement map* $B : \Sigma \rightarrow \Sigma$. The map represents what happens if all the players unilaterally and simultaneously improve their mixed strategy in response to the play σ . This map B will satisfy the requirements of Kakutani's fixed point theorem, allowing us to obtain our result.

We first define the map $B_p : \Sigma \rightarrow 2^{\Sigma_p}$ for player $p \in P$ by

$$\sigma \in \Sigma \longmapsto \operatorname{argsup}_{\sigma'_p \in \Sigma_p} \operatorname{pay}_p(\sigma - \sigma_p + \sigma'_p).$$

It is a set-valued map because the range is all the $\sigma'_p \in \Sigma_p$ at which the supremum is attained. Now define $B : \Sigma \rightarrow 2^\Sigma$ as the map that takes $\sigma \in \Sigma$ to the product set $\prod_{p \in P} B_p(\sigma)$.

We claim that a fixed point σ of B must be an equilibrium in the sense of Definition 4.2.3. This is immediate by checking that each player tolerates σ , which follows directly from the way B was defined, and from the definition of fixed point. We now show that the conditions of Kakutani's fixed point theorem (Theorem 4.4) are met by the map B .

- Since Σ_p are non-empty simplexes, the set $\Sigma = \prod_{p \in P} \Sigma_p$ of plays is convex, compact and non-empty.
- $B(\sigma)$ is non-empty for all $\sigma \in \Sigma$. This is because for each player $p \in P$, the set $B_p(\sigma)$ is non-empty, since pay_p is continuous, and the supremum of a continuous function over a compact set is always attained.
- We claim that $B(\sigma)$ is convex for all $\sigma \in \Sigma$. It is enough to argue that each set $B_p(\sigma)$ is convex since the finite product of convex sets is convex. The key observation is that, fixing σ and p , the function $f(\sigma'_p) := \operatorname{pay}_p(\sigma - \sigma_p + \sigma'_p)$ is the minimum of a finite number of linear functions, and is hence a (piecewise-linear) concave function. In other words, it is straightforward to verify that:

$$\operatorname{pay}_p(\sigma - \sigma_p + (\sigma_p^1 + \sigma_p^2)/2) \geq \frac{1}{2} (\operatorname{pay}_p(\sigma - \sigma_p + \sigma_p^1) + \operatorname{pay}_p(\sigma - \sigma_p + \sigma_p^2))$$

where $\sigma_p^1, \sigma_p^2 \in \Sigma_p$. If $\sigma_p^1, \sigma_p^2 \in B_p(\sigma)$, then $(\sigma_p^1 + \sigma_p^2)/2 \in B_p(\sigma)$ because

$$\operatorname{pay}_p(\sigma - \sigma_p + \sigma_p^1) = \operatorname{pay}_p(\sigma - \sigma_p + \sigma_p^2) = \sup_{\sigma'_p \in \Sigma_p} \operatorname{pay}_p(\sigma - \sigma_p + \sigma'_p),$$

and our claim is proved.

- We claim that the graph $\{(\sigma, B(\sigma)) \mid \sigma \in \Sigma\}$ of B is closed. This will essentially follow from the continuity of the payoff function. Consider a sequence $(\sigma(n), \tau(n))_{n \in \mathbb{N}}$ such that $\tau(n) \in B(\sigma(n))$ for all $n \in \mathbb{N}$. Suppose this sequence converges to (σ^*, τ^*) . Then continuity of pay implies that $\text{pay}(\tau(n)) \rightarrow \text{pay}(\tau^*)$.

Suppose for contradiction that $\tau^* \notin B(\sigma^*)$. Then there exists a player $p \in P$ with a mixed strategy $\tau'_p \in \Sigma_p$ such that:

$$\text{pay}_p(\tau^* - \tau_p^* + \tau'_p) > \text{pay}_p(\tau^*) = \lim_{n \rightarrow \infty} \text{pay}_p(\tau(n)).$$

We conclude that $\text{pay}_p(\tau(n) - \tau_p(n) + \tau'_p) > \text{pay}_p(\tau(n))$ for large n , because $\text{pay}_p(\tau(n) - \tau_p(n) + \tau'_p)$ can be made arbitrarily close to $\text{pay}_p(\tau^* - \tau_p^* + \tau'_p)$ by continuity of pay_p , and $\text{pay}_p(\tau(n))$ can be made arbitrarily close to $\text{pay}_p(\tau^*)$. This is a contradiction, since we had supposed that $\tau(n) \in B(\sigma(n))$. Our claim is proved.

This completes the proof. \square

4.2 The second proof

II for Theorem 4.2. Fix a min- k -game with players P . We will construct a finite game with players $P \cup \hat{P}$ where $|\hat{P}| = |P|$ and $P \cap \hat{P} = \emptyset$, with the property that each Nash equilibrium in this finite game corresponds to an equilibrium for the min-game. Since Nash equilibria exist for finite games, the theorem follows.

We now describe the finite game. For each player $p \in P$, introduce a new and distinct player $\hat{p} \in \hat{P}$. We call \hat{p} the **world-chooser opponent** of p . The set $S_{\hat{p}}$ of strategies of \hat{p} equals the set W of worlds.

We now describe the payoffs. Fix a player $p \in P$. Then the payoff $u_p : \prod_{q \in P \cup \hat{P}} S_q \rightarrow \mathbb{R}$ will depend only on $S_{\hat{p}} \times \prod_{q \in P} S_q$, and ignore the strategies played by the world-chooser opponents of the other players. If $w \in S_{\hat{p}} = W$ and $s \in \prod_{q \in P} S_q$ then we define $u_p(w, s) := u_p^w(s)$. The payoff $u_{\hat{p}}$ also depends only on the strategy played by the players P , and the player \hat{p} , and is given by $u_{\hat{p}}(w, s) = -u_p(w, s) = -u_p^w(s)$.

In words, the payoff to player p comes from the payoff function in the world chosen by \hat{p} . The world-chooser opponent \hat{p} gets a payoff that is the negative of the payoff of p . The effect of this is that player \hat{p} wants to minimize p 's payoff, by choosing the world in which p 's payoff is minimum. It remains to verify that a Nash equilibrium in this finite game corresponds to our notion of equilibrium for a min-game.

Consider a play σ . Define the **lowered** play $\ell(\sigma) := \prod_{p \in P} \sigma_p$. Then

$$\text{pay}_p(\ell(\sigma)) \leq \text{pay}_p(\sigma) \tag{1}$$

for each player $p \in P$. Here $\text{pay}_p(\ell(\sigma))$ and $\text{pay}_p(\sigma)$ denote the payoff to player p from the *min-game play* and the *finite game play* respectively. Equality holds for all plays σ tolerated by \hat{p} , in particular for all Nash equilibria.

If σ is a Nash equilibrium then we claim that $\ell(\sigma)$ is an equilibrium for the original min-game. For suppose player p is considering another mixed strategy σ'_p . We have:

$$\text{pay}_p(\ell(\sigma) - \sigma_p + \sigma'_p) = \text{pay}_p(\ell(\sigma - \sigma_p + \sigma'_p)) \leq \text{pay}_p(\sigma - \sigma_p + \sigma'_p) \leq \text{pay}_p(\sigma) = \text{pay}_p(\ell(\sigma))$$

The first inequality is by Equation 1. The last equality is by Equation 1, using the fact that σ is a Nash equilibrium. Hence every player tolerates $\ell(\sigma)$, and so it is an equilibrium by definition. \square

5 Reducing Kripke games to min-games

Theorem 5.1. Every Kripke game has an equilibrium play.

Proof. We will deduce this theorem by reducing a Kripke game to a min-game with the property that the min-game's equilibria correspond to equilibria of the original Kripke game. The result then follows from Theorem 4.2.

Fix a Kripke game with players P on worlds W . We will construct a min- $|W|$ -game with players

$$\mathcal{P} = \cup_{p \in P} [W]_p.$$

That is, each player of the original Kripke game is split into multiple players, one for each \equiv_p -equivalence class. Player $[w]_p$ is assigned the strategy set S_p of player p in the original Kripke game. On world $v \in W$, the payoff

$$u_{[w]_p}^v \left(\prod_{p \in \mathcal{P}} s_p \right) = \begin{cases} u_p^v \left(\prod_{q \in P} s_{[w]_q} \right) & \text{if } v \equiv_p w \\ +A & \text{otherwise.} \end{cases}$$

where u_p^v is the payoff from the original Kripke game, and $A > 0$ is a positive constant chosen sufficiently large so that every player's minimum payoff is guaranteed to be less than A , in every play.

In words, we have ensured that the payoff to player $[w]_p$ comes only from worlds $v \equiv_p w$. The only opponents who matter to $[w]_p$ are players $[w]_q$ where $q \in P$.

Given a play σ for this min-game, we construct a play τ for the original Kripke game as follows: the mixed strategy $\tau_p^{[w]_p}$ is defined as $\sigma_{[w]_p}$. The key property is that

$$\text{pay}_p^w(\tau) = \text{pay}_{[w]_p}(\sigma)$$

where the LHS is the payoff for the Kripke game, and the RHS for the min-game. This allows us to prove that if σ is an equilibrium play for this min-game then τ is an equilibrium play for the original Kripke game. If not, then player p can improve her payoff in world w in the Kripke game, which will translate to an improved payoff for player $[w]_p$ in the min-game, contradicting the assumption that σ was an equilibrium play. \square

6 Related work

Games with incomplete information, and epistemic notions in game theory, have appeared before, in particular in several works by Aumann. The closest to our work appears to be Aumann's work on Interactive Epistemology [1]. There are parallels between Aumann's *semantic approach* and our approach. Aumann's space Ω of states of the world appears to correspond to our set W of worlds. His partitions on the world appear to correspond to the equivalence class partition induced by the equivalence relations \equiv_p for each player p . The focus in [1] appears to be to provide epistemic justifications for many of the standard assumptions of game theory, whereas our focus is on how epistemic considerations affect strategy.

References

- [1] Robert J. Aumann, *Interactive epistemology i: Knowledge*, International Journal of Game Theory (1999).
- [2] Daniel Kahneman and Amos Tversky, *Prospect theory: An analysis of decision under risk*, Econometrica **47** (1979), no. 2, 263–91.
- [3] Saul A. Kripke, *A completeness theorem in modal logic*, The Journal of Symbolic Logic **24** (1959), no. 1, pp. 1–14 (English).
- [4] K. Miettinen, *Nonlinear multiobjective optimization*, Kluwer Academic Publishers, Boston, 1999.
- [5] John Von Neumann and Oskar Morgenstern, *Theory of games and economic behavior (commemorative edition)*, Princeton university press, 2007.
- [6] M.J. Osborne and A. Rubinstein, *A course in game theory*, MIT Press, 1994.